



ELSEVIER

Available online at www.sciencedirect.com



ScienceDirect

Journal of Computational and Applied Mathematics 223 (2009) 735–752

JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICS

www.elsevier.com/locate/cam

The finite element-Galerkin method for singular self-adjoint differential equations

Mohamed A. El-Gebeily^{a,*}, Khaled M. Furati^a, Donal O'Regan^b

^a *Mathematical Sciences Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia*

^b *Department of Mathematics, National University of Ireland, Galway, Ireland*

Received 12 November 2005; received in revised form 4 December 2007

Abstract

We investigate the finite element-Galerkin method for singular self-adjoint second-order differential expressions. The weak formulation of the problem involves integration by parts, which allows the use of the usual piecewise linear functions. Our analysis shows that the method produces the solution corresponding to a particular self-adjoint realization of the differential expression. We also propose two algorithms to approximate the solution of any self-adjoint realization. Numerical examples are given to illustrate the analysis as well as the proposed algorithms.

© 2008 Elsevier B.V. All rights reserved.

Keywords: Singular differential operators; Self-adjoint operators; Weak formulation; Finite element method; Galerkin method

1. Introduction

We investigate the finite element-Galerkin method for the differential equations generated by the self-adjoint expression

$$\ell u = \frac{1}{w} \left(- (pu')' + qu \right).$$

Our main goal is to treat the singular case although the regular case can also be handled in the same way. The advantage of the Galerkin method is that we can do integration by parts and this reduces the smoothness requirements of the basis functions. This is particularly important in the case at hand because it will relieve us from considering the effects of the properties of the coefficient functions on such smoothness. For example, in order for the expression ℓ to make sense, we need pu' to be absolutely continuous. Thus the properties of p need to be accounted for. On the other hand, by integrating by parts we deal with the sesquilinear form

$$a(u, v) = \int pu' \bar{v}' + qu \bar{v} \tag{1}$$

in which no such smoothness is required.

* Corresponding author. Tel.: + 966 3 8603728; fax: + 966 3 8602340.

E-mail address: mgebeily@kfupm.edu.sa (M.A. El-Gebeily).

Symmetric sesquilinear forms associated with the expression ℓ in the singular case take the form $\langle \ell u, v \rangle (= \langle u, \ell v \rangle)$, where $\langle \cdot, \cdot \rangle$ is the inner product in a Hilbert space (see, for example, [11,13,21]). This form takes two integrations by parts in order to move the expression from one function to the other and has the disadvantages discussed above. The symmetric form $a(\cdot, \cdot)$ was dealt with more recently in [6,7] including discussion of the boundary conditions associated with it.

The variational method was used, in connection with self-adjoint operators by many authors for the calculation of guaranteed bounds for the k lowest eigenvalues [2,18,22], approximating the eigenvalues [14] analyzing the asymptotic behavior of the errors [15], studying the spurious eigenvalues generated by the method in a gap between two parts of the essential spectrum [4,12]. On the other hand, it was stated in [19] that the variational method has lost its glamour as a practical method for approximating the eigenvalues of Sturm–Liouville problems. In this work we hope to establish that, at least for the direct problem, the method is still very vital for handling the general singular expressions. We show here that the finite element implementation of the variational method is capable of approximating the solution of such singular expressions under all classifications of the problem (see the next section).

The analysis carried out here shows that the approximations obtained by the method converge to the solution of a particular self-adjoint realization of the problem. Methods for finding the solution to an arbitrary self-adjoint realization are then developed based on the boundary condition characterization of a given realization. Specifically, we propose two algorithms for approximating the solution corresponding to a general self-adjoint realization of the expression.

This paper consists of seven sections in addition to the introduction. Section 2 contains some preliminary theory and terminology specific to the singular self-adjoint differential operators. Section 3 reviews the characterization of self-adjoint differential operators in terms of “boundary conditions”. In Section 4 we set the variational formulation of the problem and show its equivalence to a particular self-adjoint operator. Section 5 states the discretization of the problem. Section 6 discusses the convergence of the approximate solutions to that of the original problem which is equivalent to the variational formulation. In Section 7 we develop two algorithms to approximate the solution corresponding to an arbitrary self-adjoint operator. Numerical examples are given in Section 8 that illustrate the analysis as well as the algorithms.

2. Preliminaries

In this section we introduce notation, definitions and known results necessary for this work. The presentation in this section is taken from [10,21,16,20]. We work with the formally self-adjoint differential expression

$$\ell u = \frac{1}{w} \left[(-pu')' + qu \right]$$

defined on the interval $I = (a, b)$, $-\infty \leq a < b \leq \infty$. We assume that

$$1/p, q, w \in L_{\text{loc}}(I),$$

p and q are real-valued and that $w > 0$ almost everywhere in I .

Let $H = L_w^2(I)$, be the Hilbert space of square integrable functions with respect to the weight w . The inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$ in this space are given by

$$\langle f, g \rangle = \int_I f(t) \bar{g}(t) w(t) dt$$

and

$$\|f\|^2 = \int_I |f(t)|^2 w(t) dt,$$

respectively. Also let $u^{[1]} = pu'.u^{[1]}$ be called the first pseudoderivative of u with respect to the functions p, q . The maximal operator L generated by the expression ℓ in H is defined by

$$D(L) = D = \left\{ u \in H : u, u^{[1]} \in AC(I) \text{ and } \ell u \in H \right\}, \quad Lu = \ell u, \quad u \in D.$$

Since D is dense in H , it has a uniquely defined adjoint. Let $L_0 = L^*$ and $D_0 = D(L_0)$. The operator L_0 is called the minimal operator generated by ℓ and it is known [16] that $D_0 \subseteq D$, D_0 is dense in H and $L_0^* = L$. In other words, $L_0 \subset L = L_0^*$. Therefore, L_0 is a symmetric closed operator. Moreover, any self-adjoint extension of L_0 is a self-adjoint restriction of L and vice versa, i.e., $L_0 \subset S = S^* \subset L_0^* = L$. The pre-minimal operator L'_0 is defined by

$$D'_0 = D(L'_0) = \{u \in D_0 : \text{supp } u \subset I\}, \\ L'_0 u = L_0 u \quad \forall u \in D'_0.$$

L_0 is the closure of the operator L'_0 .

For a fixed non-real λ , let R_λ denote the range of $L_0 - \lambda E$, where E is the identity operator on H . The deficiency space N_λ of L_0 is defined to be the orthogonal complement of R_λ in H , i.e.,

$$N_\lambda = R_\lambda^\perp.$$

It is shown in [16] that

$$N_\lambda = \{y \in H : L_0^* y = L y = \bar{\lambda} y\}, \\ D = D_0 \dot{+} N_\lambda \dot{+} N_{\bar{\lambda}},$$

and

$$\dim(N_\lambda) = \dim(N_{\bar{\lambda}}).$$

Here, $X \dot{+} Y$ means the direct sum of the two not necessarily orthogonal subspaces X and Y . We denote the common value, $\dim(N_\lambda)$, by d and call d the *deficiency index* of L_0 on I . Furthermore, if $\lambda \in \mathbb{R}$ is a point of regular type (see [9]) for L_0 , then $d = \dim(N_\lambda)$. It is shown in [16] that $0 \leq d \leq 2$, and if one endpoint is regular and the other is singular (see definitions below), then $1 \leq d \leq 2$. Hence, D is a $2d$ -dimensional extension of D_0 .

For $y, z \in D$ and $x \in I$ define the Lagrange bracket

$$[y, z](x) = y(x) \bar{z}^{[1]}(x) - \bar{z}(x) y^{[1]}(x). \quad (2)$$

Note that the limits of the terms in (2) as $x \rightarrow a^+, b^-$ both exist and are finite. Thus, the notation

$$[y, z](a) = \lim_{x \rightarrow a^+} [y, z](x), \quad [y, z](b) = \lim_{x \rightarrow b^-} [y, z](x)$$

is justified. We use $[y, z]_\alpha^\beta$ to denote $[y, z](\beta) - [y, z](\alpha)$.

Proposition 1. *The following relation holds between D_0 and D in general*

$$D_0 = \{y \in D : [y, z]_a^b = 0, \forall z \in D\}. \quad (3)$$

Proof. See [16]. ■

The (finite or infinite) endpoint a is called regular if $1/p, q, w \in L(a, c)$ for some (and hence all) $c \in I$; is limit circle (LC) if all solutions of

$$\ell u = \lambda u$$

are in $L_w^2(a, c)$ for some $\lambda \in \mathbb{C}$ and $c \in I$; is limit point (LP) if it is not LC. Similar definitions hold at b . An endpoint is singular if it is not regular. The LC and LP classifications are independent of $\lambda \in \mathbb{C}$. See [21,10] for more details.

Proposition 2. 1. $d = 0 \iff a$ and b are LP.
2. $d = 1 \iff$ one endpoint is LP and the other is LC.
3. $d = 2 \iff a$ and b are LC.

Proof. See [16], page 72. ■

The proofs of the following two lemmas can be found in [8,21].

Lemma 3. Suppose a (b) is LC, then there are real ψ_1, ψ_2 (ψ_3, ψ_4) $\in D \setminus D_0$ such that

1. $[\psi_1, \psi_2](a) \neq 0$ ($[\psi_3, \psi_4](b) \neq 0$).
2. ψ_1 and $\psi_2 = 0$ near b (ψ_3 and $\psi_4 = 0$ near a).
3. ψ_1 and ψ_2 (ψ_3 and ψ_4) are linearly independent modulo D_0 .

We remark that in the above lemma, if both a and b are LC, then $\{\psi_1, \psi_2, \psi_3, \psi_4\}$ can be chosen so that they are linearly independent modulo D_0 .

Lemma 4. Suppose a (b) is LP, then there is a real ψ_a (ψ_b) $\in D \setminus D_0$ such that $\psi_a = 0$ near b ($\psi_b = 0$ near a).

Let $c \in I$ and θ, ϕ be the unique (real) solutions of the initial value problems

$$\ell u = 0, \quad (4)$$

$$\theta(c) = -\phi^{[1]}(c) = 1, \quad (5)$$

$$\theta^{[1]}(c) = \phi(c) = 0. \quad (6)$$

If a is LC then θ, ϕ belong to $L_w^2(a, c)$. In this case ψ_1, ψ_2 may be constructed by taking them equal to θ, ϕ , respectively, near a and equal to 0 near b . If a is LP then for some real m , the linear combination $\psi_a = \theta + m\phi$ belongs to $L_w^2(a, c)$. Similar comments hold for the endpoint b .

Lemma 5. If a and b are both LP, then

$$D = D_0.$$

If a is LC and b is LP, then

$$D = D_0 \dot{+} \text{span}\{\psi_1, \psi_2\}.$$

If b is LC and a is LP, then

$$D = D_0 \dot{+} \text{span}\{\psi_3, \psi_4\}.$$

If a and b are both LC then

$$D = D_0 \dot{+} \text{span}\{\psi_1, \psi_2, \psi_3, \psi_4\},$$

where $\psi_1, \psi_2, \psi_3, \psi_4$ are as in Lemma 3.

Any self-adjoint realization \widehat{L} of the formal operator ℓ in H is an extension of L_0 (restriction of L): $L_0 \subset \widehat{L} \subset L$. The domain of definition \widehat{D} of such a realization is characterized by the existence of d functions ξ_1, \dots, ξ_d (the list is empty if $d = 0$) in D which are linearly independent modulo D_0 and satisfy $[\xi_i, \xi_j]_a^b = 0$ such that

$$\widehat{D} = D_0 \dot{+} \text{span}\{\xi_1, \dots, \xi_d\}.$$

Next we introduce the sesquilinear form

$$a(u, v) = \int_a^b p u' \bar{v}' + q u \bar{v}$$

and the associated boundary terms

$$\begin{aligned} \{u, v\}(x) &= -u^{[1]} \bar{v}(x), \quad x \in I, \\ \{u, v\}_a^b &= \{u, v\}(b^+) - \{u, v\}(a^-) \end{aligned}$$

whenever the implied limits exist. Note that

$$[u, v](x) = \{u, v\}(x) - \{\bar{v}, \bar{u}\}(x).$$

Also, for $u, v \in D$, $a(u, v)$ exists and is finite if and only if $\{u, v\}_a^b$ exists and is finite. Then

$$a(u, v) = \langle Lu, v \rangle - \{u, v\}_a^b. \quad (7)$$

The following set is defined in connection with the form (1):

$$W = \{u \in H : |a(u, u)| < \infty\}.$$

Note that the space W contains the space D'_0 . Hence, W is dense in H .

Assume that $a(\cdot, \cdot)$ is elliptic on W : For some $\mu > 0$,

$$a(u, u) \geq \mu \|u\|^2 \quad \forall u \in W. \quad (8)$$

It follows from the Cauchy–Schwartz inequality that $|a(u, v)| < \infty$ for all $u, v \in W$ and, therefore, W is a subspace of H . Obviously, W is the maximal subspace of H on which $a(\cdot, \cdot)$ can be defined. The W ellipticity of $a(\cdot, \cdot)$ means that the mappings $(u, v) \mapsto a(u, v)$ and $u \mapsto \sqrt{a(u, u)}$ define an inner product and its corresponding norm on W and that, equipped with this norm, W is continuously embedded in H . We observe that W is a Hilbert space in this norm. Indeed, if we denote by \overline{W} the closure of W in this norm, then the continuous embedding of W in H extends to a continuous embedding of \overline{W} in H . In other words, \overline{W} is a subspace of H with $a(u, u) < \infty$ for all $u \in \overline{W}$. Therefore, $\overline{W} \subseteq W$.

In what follows, we will denote the norm and inner product produced by $a(\cdot, \cdot)$ on W by $\|\cdot\|_W$ and $\langle \cdot, \cdot \rangle_W$, respectively.

3. Self-adjoint boundary conditions

In this section we characterize the self-adjoint extensions of L_0 by their boundary conditions. For a given $u \in D$ we will use the notation

$$U(x) = \begin{bmatrix} [u, \theta](x) \\ [u, \phi](x) \end{bmatrix}, \quad x \in I$$

and $U(a)(U(b))$ to mean $U(a^+)(U(b^-))$. As mentioned in the preliminaries, for $u, v \in D$, $[u, v](a)$ and $[u, v](b)$ are finite. Moreover, it was shown in [10] that the values of these brackets are arbitrary complex numbers and, given any pair of complex numbers, one can find a pair of functions $u, v \in D$ whose bracket values match the given complex numbers. Therefore, naturally, when boundary conditions are to be considered for singular problems, these brackets replace the ordinary function values typically considered in the regular case.

We first observe that, in the regular case ($1/p, q, w$ are integrable on I), the bracket boundary conditions are equivalent to function value conditions. To see this let θ, ϕ be the solutions of (4)–(6). It is known that (see [16] page 54) $D \subset AC(\overline{I})$. The equation

$$\begin{bmatrix} [u, \theta](a) \\ [u, \phi](a) \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (9)$$

is equivalent to

$$\begin{bmatrix} \theta^{[1]}(a) & -\theta(a) \\ \phi^{[1]}(a) & -\phi(a) \end{bmatrix} \begin{bmatrix} u(a) \\ u^{[1]}(a) \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \quad (10)$$

On the other hand, the equations

$$\int_c^x \ell u \bar{v} w \, dt = [u, v]_c^x + \int_c^x u \bar{\ell} \bar{v} w \, dt$$

and $\ell\theta(t) = \ell\phi(t) = 0$ imply that $[\theta, \phi]_c^x = 0$ for all $x \in I$. Therefore, $[\theta, \phi](x) = [\theta, \phi](c) = -1$. In particular, we have $[\theta, \phi](a) = -1$. This is precisely the determinant of system (10). In other words this system has a unique solution. Thus assuming boundary conditions of the form (9) is equivalent to assuming classical boundary conditions.

The boundary conditions assumed by self-adjoint extensions of L_0 will depend on the classification of each endpoint. If $d = 0$ then each endpoint is LP. In this case L_0 is self-adjoint and no boundary conditions are allowed. If $d = 1$ then one endpoint is regular or LC and the other point is LP. Assume without loss of generality that a is regular or LC and b is LP.

Lemma 6. Assume that $d = 1$, a is regular or LC and b is LP. Let $\psi_1, \psi_2 \in D$ be real-valued such that $[\psi_1, \psi_2](a) = 1$. A symmetric extension \widehat{L} of L_0 with domain \widehat{D} is self-adjoint if and only if there exists a $\gamma \in [0, \pi)$ such that

$$\widehat{D} = \{u \in D : CU(a) = 0\}, \quad (11)$$

where $C = [\cos \gamma \quad \sin \gamma]$ and

$$U(a) = \begin{bmatrix} [u, \psi_1](a) \\ [u, \psi_2](a) \end{bmatrix}.$$

Proof. See [8]. ■

If $d = 2$ then both endpoints are either regular or LC. In this case we can find two real-valued functions $\psi_1, \psi_2 \in D$ such that $[\psi_1, \psi_2](a) \neq 0$ and $[\psi_1, \psi_2](b) \neq 0$ (e.g., we can take $\psi_1 = \theta, \psi_2 = \phi$). These two conditions imply that ψ_1, ψ_2 are linearly independent modulo D_0 . Indeed if $r\psi_1 + s\psi_2 \in D_0$, then

$$\begin{aligned} 0 &= [\psi_1, r\psi_1 + s\psi_2](a) \\ &= r[\psi_1, \psi_1](a) + s[\psi_1, \psi_2](a) \\ &= s[\psi_1, \psi_2](a). \end{aligned}$$

Thus $s = 0$. Also

$$\begin{aligned} 0 &= [\psi_2, r\psi_1 + s\psi_2](b) \\ &= r[\psi_2, \psi_1](b) \end{aligned}$$

gives $r = 0$. Therefore, by replacing ψ_1, ψ_2 by appropriate linear combinations of them, if necessary, we may normalize to $[\psi_1, \psi_2](a) = 1 = [\psi_1, \psi_2](b)$.

Lemma 7. Assume that $d = 2$. Let $\psi_1, \psi_2 \in D$ be real-valued such that $[\psi_1, \psi_2](a) = 1 = [\psi_1, \psi_2](b)$. A symmetric extension \widehat{L} of L_0 with domain \widehat{D} is self-adjoint if and only if one of the following two conditions holds

1. (Separated boundary conditions.) There exist $\gamma_1, \gamma_2 \in [0, 2\pi)$ and $r_1, r_2 \in [0, 1]$ such that

$$\widehat{D} = \{u \in D : C_a U(a) = C_b U(b) = 0\},$$

where

$$\begin{aligned} U(a) &= \begin{bmatrix} [u, \psi_1](a) \\ [u, \psi_2](a) \end{bmatrix}, \quad U(b) = \begin{bmatrix} [u, \psi_1](b) \\ [u, \psi_2](b) \end{bmatrix}, \\ C_a &= \begin{bmatrix} r_1 e^{i\gamma_1} & \sqrt{1-r_1^2} \end{bmatrix}, \quad C_b = \begin{bmatrix} r_2 e^{i\gamma_2} & \sqrt{1-r_2^2} \end{bmatrix}. \end{aligned}$$

2. (Coupled boundary conditions.) There exists a real 2×2 matrix C with $\det C = 1$ and a $\gamma \in [0, 2\pi)$ such that

$$\widehat{D} = \left\{ u \in D : U(a) = e^{i\gamma} C U(b) \right\}. \quad (12)$$

Proof. See [21]. ■

We remark that the functions ψ_1, ψ_2 are also known as the regularizing functions. We also remark that if either endpoint is regular, then in the above two lemmas the boundary conditions stated in terms of the Lagrange brackets at the endpoints can be replaced with boundary conditions stated in terms of the function values. See the discussion at the beginning of this section.

4. Weak formulation

In this section we study the weak formulation of the direct problem $\ell u = f$ in the space W together with the self-adjoint operators induced by $a(\cdot, \cdot)$ in the space H .

Proposition 8. $D_0 \subseteq W$ and, for all $u \in D_0$ and $v \in W$, $a(u, v) = \langle L_0 u, v \rangle$. Consequently $\{u, v\}_a^b = 0$ for all $u, v \in D_0$.

Proof. Let $u \in D_0$. Since L_0 is the closure of the symmetric operator L'_0 we can find a sequence $\{u_n\}_{n=1}^\infty$ in D'_0 such that $u_n \rightarrow u$ and $L'_0 u_n = L_0 u_n \rightarrow L_0 u$ (both in H). Then

$$\begin{aligned} |a(u_n - u_m, u_n - u_m)| &= |\langle L_0(u_n - u_m), u_n - u_m \rangle| \\ &\leq \|L_0(u_n - u_m)\| \|u_n - u_m\|. \end{aligned}$$

From this it follows that $\{u_n\}_{n=1}^\infty$ is a Cauchy sequence in W . Therefore, $u_n \rightarrow u$ in W . To show the second statement let $u \in D_0$, $v \in W$ and let $\{u_n\}_{n=1}^\infty$ be a sequence in D'_0 such that $u_n \rightarrow u$ and $L_0 u_n \rightarrow L_0 u$ (both in H). We argue as before that $u_n \rightarrow u$ in W . Then

$$a(u, v) = \lim a(u_n, v) = \lim \langle L_0 u_n, v \rangle = \langle L_0 u, v \rangle.$$

Finally, if $u, v \in D_0$ then

$$a(u, v) = \langle L_0 u, v \rangle = \langle u, L_0 v \rangle.$$

Hence $\{u, v\}_a^b = 0$ for all $u, v \in D_0$. ■

Observe that the above argument actually shows that D'_0 is dense in D_0 in the topology of W . Also the equation $a(u, v) = \langle L_0 u, v \rangle$ for all $u \in D_0$, $v \in W$ means that, for a fixed $u \in D_0$, $a(u, \cdot)$ is continuous on D_0 with respect to the norm in H . This prompts us to define the set

$$\tilde{D} = \{u \in W : a(u, \cdot) \text{ is continuous on } D_0 \text{ with respect to the norm in } H\}. \quad (13)$$

The next proposition establishes some properties of functions in \tilde{D} and, in particular, the fact that for $d \geq 1$, \tilde{D} is an essential extension of D_0 .

Proposition 9. Let \tilde{D} be defined by (13). Then

1. $\tilde{D} \subset D$ and, for all $u \in \tilde{D}$ and $v \in D_0$, $\{u, v\}_a^b = \{v, u\}_a^b = 0$,
2. for $u \in \tilde{D}$ and $v \in W$ both $\{u, v\}$ (a) and $\{u, v\}$ (b) exist and are finite,
3. $\tilde{D} = \{u \in D : \{u, v\}_a^b = 0 \forall v \in D_0\}$ and,
4. for $d \geq 1$, there are d functions in \tilde{D} that are linearly independent modulo D_0 .

Proof. To show 1, let $u \in \tilde{D}$. For any $v \in D_0$, $\langle u, L_0 v \rangle = a(u, v)$ by Proposition 8. Since $a(u, \cdot)$ is continuous on D_0 by the definition of \tilde{D} , the functional $v \mapsto \langle u, L_0 v \rangle$ is continuous on D_0 . Therefore, $u \in D$ and $a(u, v) = \langle u, L_0 v \rangle = \langle Lu, v \rangle$. Thus $\{u, v\}_a^b = \{v, u\}_a^b = 0$ for all $u \in \tilde{D}$ and $v \in D_0$.

To show 2, take x, y such that $a < x < y < b$. Then

$$\int_x^y pu'v' + qu\bar{v} = -\{u, v\}(y) + \{u, v\}(x) + \int_x^y \ell u\bar{v}w.$$

Fixing x and taking the limit on both sides as $y \rightarrow b^-$, we see that $\{u, v\}$ (b) is finite. Similarly we see that $\{u, v\}$ (a) is finite.

To show 3, it suffices in light of property 1, to show that $\{u \in D : \{u, v\}_a^b = 0 \forall v \in D_0\} \subset \tilde{D}$. So let $u \in D$ with $\{u, v\}_a^b = 0 \forall v \in D_0$. Then, for any $v \in D_0$, $\langle Lu, v \rangle$ exists and by Eq. (7), since $\{u, v\}_a^b = 0$, $a(u, v)$ exists and $a(u, v) = \langle Lu, v \rangle$. From this we get that $u \in \tilde{D}$.

To show 4, notice that since $d = \dim(R(L_0)^\perp) = \ker(L)$, we can choose d linearly independent functions $f_1, \dots, f_d \in \ker(L)$. For $k = 1, \dots, d$ the variational problem

$$a(u, v) = \langle f_k, v \rangle \quad \forall v \in W \quad (14)$$

has a unique solution u_k by the Lax–Milgram Lemma. These solutions are in \tilde{D} but cannot be in D_0 for otherwise, using Proposition 8, we can show that $L_0 u_k = f_k$ which contradicts the choice of f_k . The linear independence of f_1, \dots, f_d and their orthogonality to $R(L_0)$ imply that u_1, \dots, u_d are linearly independent modulo D_0 . ■

Proposition 10. Let $u_k \in \tilde{D}$ be the solution of $Lu = f_k, k = 1, \dots, d$ where the functions $f_1, \dots, f_d \in \ker(L)$ are linearly independent. Then $\{u_i, u_j\}_a^b = 0, i, j = 1, \dots, d$.

Proof. Since $u_i \in D, i = 1, \dots, d$,

$$\langle Lu_i, v \rangle = a(u_i, v) = \langle f_i, v \rangle$$

for any $v \in D'_0$. Hence, $Lu_i = f_i$. Now, for all $v \in W$,

$$\langle Lu_i, v \rangle = \langle f_i, v \rangle = a(u_i, v).$$

Therefore, $\{u_i, v\}_a^b = 0$. In particular, $\{u_i, u_j\}_a^b = 0, i, j = 1, \dots, d$. ■

Corollary 11. Let u_1, \dots, u_d be as in the previous proposition. Define the domain

$$D_1 = D_0 \dot{+} \text{span}\{u_1, \dots, u_d\}.$$

Then D_1 is the domain of definition of a self-adjoint extension L_1 of L_0 .

Proof. Since

$$\langle Lu, v \rangle = a(u, v) = \langle u, Lv \rangle \quad \forall u, v \in D_1,$$

the restriction of L to D_1 is symmetric. Since D_1 is a d -dimensional extension of D_0 , it follows from the characterizations of the domains of self-adjoint extensions of L_0 (restrictions of L) given in Naimark [16] that this restriction is self-adjoint. ■

Definition 12. Self-adjoint extensions of L_0 corresponding to the various choices of the functions $u_1, \dots, u_d \in \tilde{D}$ which are linearly independent modulo D_0 and satisfy the boundary conditions $\{u_i, u_j\}_a^b = 0, i, j = 1, \dots, d$ will be called type I operators.

It should be noted here that not all self-adjoint extensions of L_0 are type I operators. For example, consider the expression $\ell u = -u'' + u$ defined on the interval $I = (0, 1)$. The boundary conditions $u(0) + u'(0) = u(1) + u'(1) = 0$ give rise to a self-adjoint extension \hat{L} of L_0 in the space $L^2(I)$. The function $u(x) = -3x^3 + 4x^2$ is in the domain of this operator but $\{u, u\}_0^1 \neq 0$. We would like to comment also about type I operators and the Friedrichs extension [5, 17]. In the regular case, it is known that the latter one satisfies Dirichlet boundary conditions while type I operators satisfy more general (separated as well as nonseparated) boundary conditions. Hence, the Friedrichs extension is a type I operator. In the singular case, the domain D_F of the Friedrichs extension is characterized by the Lagrange boundary conditions $\{u, v\}_a^b = 0$ for all $u, v \in D_F$. Since the Friedrichs extension and type I operators are defined by the same sesquilinear form, we have a sharper condition $\{u, v\}_a^b = 0$ for all $u, v \in D_F$.

Given a type I operator L_1 with domain D_1 we will call the subspace

$$W_1 = \left\{ u \in W : \{z, u\}_a^b = 0 \quad \forall z \in D_1 \right\} \quad (15)$$

the associated subspace. It immediately follows from the definition of W_1 that

$$a(u, v) = \langle L_1 u, v \rangle \quad \forall u \in D_1, v \in W_1. \quad (16)$$

The converse of this statement is also true, namely, for a given $f \in H$, the statement

$$u \in W_1 \quad \text{and} \quad a(u, v) = \langle f, v \rangle \quad \forall v \in W_1$$

implies that $u \in D_1$ and $L_1 u = f$. Indeed, for any for $v \in D_1$

$$\langle u, L_1 v \rangle = a(u, v) = \langle f, v \rangle.$$

It follows that the functional $v \mapsto \langle u, L_1 v \rangle$ is continuous on D_1 . Hence $u \in D(L_1^*) = D_1$ and $L_1 u = f$. The following proposition gives two important properties of W_1 .

Proposition 13. Suppose L_1 is a type I operator with domain D_1 and associated subspace W_1 . Then W_1 is a closed subspace of W and D_1 is dense in W_1 in the topology of W .

Proof. Let $\{v_n\}_{n=1}^\infty$ be a sequence in W_1 such that $v_n \rightarrow v$ in W (and in H). For any $z \in D_1$ we have

$$a(v, z) = \lim a(v_n, z) = \lim \langle v_n, L_1 z \rangle = \langle v, L_1 z \rangle.$$

Hence $\{z, v\}_a^b = 0$. Thus $v \in W_1$. To prove the density statement assume that $u \in W_1$ is orthogonal to D_1 . Then, for any $v \in D_1$, $a(u, v) = 0$. Therefore,

$$\langle u, L_1 v \rangle = a(u, v) = 0 \quad \forall v \in D_1.$$

Thus u is orthogonal to the range of L_1 and, hence $u \in \ker L_1$ as L_1 is self-adjoint. But since L_1 is one-to-one (by (8)), we get that $u = 0$. ■

We have thus proven the following theorem.

Theorem 14. Suppose L_1 is a type I operator with domain D_1 and associated subspace W_1 . The following statements are equivalent for a given $f \in H$.

1. $u \in D_1$ and

$$L_1 u = f.$$

2. $u \in W_1$ and

$$a(u, v) = \langle f, v \rangle \quad \forall v \in W_1. \quad (\text{V})$$

A special case of Theorem 14, which is of particular interest to us is when $f \in R_0$, the range of L_0 . In this case we can state the following corollary.

Corollary 15. Let $f \in R_0$. The following are equivalent.

1. $u \in D_0$ and $L_0 u = f$.

2. $u \in W_1$ and $a(u, v) = \langle f, v \rangle \quad \forall v \in W_1$.

We close this section with a remark on the ellipticity condition (8). This condition is weaker than assuming that a given self-adjoint extension of L_0 is positive definite. For example, consider the expression $\ell u = -u'' + u$ on $(0, \pi)$. The corresponding sesquilinear form is positive definite and the corresponding minimal operator L_0 has no eigenvalues. It is known, however, (see, for example, [9]) that given any real number λ_0 , there exists a self-adjoint extension \widehat{L} of L_0 for which λ_0 is an eigenvalue.

5. The basis functions

We construct a sequence of basis functions which will be suitable for implementing the finite element-Galerkin method. Let $\{a_N\}, \{b_N\}$ be two monotonic sequences such that $a_N \searrow a$ and $b_N \nearrow b$. It is to be understood that if $a(b)$ is regular and finite then it is included in the interval I and $a_N = a(b_N = b)$. For $n = 2, 3, \dots$, let

$$h_{n,N} = \frac{b_N - a_N}{n}.$$

The interval $I_N = [a_N, b_N]$ is divided into n equal subintervals at the nodes

$$t_{i,n,N} = a_N + i h_{n,N}, \quad i = 1, 2, \dots, n-1.$$

The function $\eta_{i,n,N}$ defined by

$$\eta_{i,n,N}(t) = \max \left\{ 1 - h_{n,N}^{-1} |t - t_{i,n,N}|, 0 \right\}, \quad t \in I$$

has the following familiar properties:

1. $\eta_{i,n,N}$ is absolutely continuous and piecewise linear.

2. $\eta_{i,n,N}(t_{j,n,N}) = \delta_{ij}$.
3. $\text{supp } \eta_{i,n,N} = [t_{i-1,n,N}, t_{i+1,n,N}] \subseteq I_N \subseteq I$.

Proposition 16. Let V be the subspace of W of functions with compact support in I , then $\eta_{i,n,N} \in V$ for all $N = 1, 2, \dots, n = 2, 3, \dots, i = 1, 2, \dots, n - 1$.

Proof. Since w is locally integrable, for any compact subinterval $[\alpha, \beta] \subset I$ and any $k \geq 0$, $\int_{\alpha}^{\beta} t^k w(t) dt < \infty$. Therefore, $\eta_{i,n,N} \in H$. Also since $\eta_{i,n,N}$ is absolutely continuous, it has an integrable derivative and compact support in I , then $\eta_{i,n,N} \in V$. ■

Let $\{w_k\}_{k=1}^{\infty}$ be a remuneration of the set

$$\{\eta_{i,n,N} : i = 1, 2, \dots, n - 1, n = 2, 3, \dots, N = 1, 2, \dots\}.$$

and let $V_m = \text{span}\{w_1, w_2, \dots, w_m\}$, $m = 1, 2, \dots$. It follows from the above proposition that, for all m , $V_m \subset W$. Define the dense subspace S of H by

$$S = \bigcup_{m=1}^{\infty} V_m.$$

Notice that if the end point a (b) is regular and finite then for any $v \in S$, $v(a) = 0$ ($v(b) = 0$); furthermore, for any $z \in D_1$, $z(a)$ and $z^{[1]}(a)$ ($z(b)$ and $z^{[1]}(b)$) exist and are finite because this is true [16] about any function in D .

Proposition 17. Suppose L_1 is a type I operator with domain D_1 and associated subspace W_1 . If the function p is essentially locally bounded and $1/p$ is locally square integrable then the space S is dense in W_1 .

Proof. Define the subspace

$$W_2 = \left\{ u \in W_1 : \{u, z\}_a^b = 0 \ \forall z \in D_1 \right\}. \quad (17)$$

Observe that W_2 is dense in W_1 since $D_1 \subset W_2 \subset W_1$. Therefore, it suffices to show that S is dense in W_2 . It follows from (15), (17) and the observations prior to this proposition that $S \subset W_2$. Now suppose $u \in W_2$ is orthogonal to S in the topology of W . Then for any $v \in S$,

$$a(u, v) = 0. \quad (18)$$

We claim that Eq. (18) is true also for any $v \in D_0$. Since D'_0 is dense in D_0 in the topology of W , given $v \in D_0$ we can choose a $v_1 \in D'_0$ such that $\|v - v_1\|_W < \varepsilon$. We proceed to show that we can choose a $v_2 \in S$ such that $\|v_2 - v_1\|_W < \varepsilon$. Let $\text{supp } (v_1) = [\alpha, \beta] \subset (a, b)$. First, observe that $v_1 \in H_0^1[\alpha, \beta]$. To see this we observe first that v_1 is absolutely continuous on $[\alpha, \beta]$. Then, because of the local square integrability of $1/p$, we have

$$\int_{\alpha}^{\beta} (v'_1)^2 = \int_{\alpha}^{\beta} \frac{1}{p^2} (pv'_1)^2 \leq \|(pv'_1)^2\|_{\infty} \int_{\alpha}^{\beta} \frac{1}{p^2}.$$

Since $S|_{(\alpha, \beta)}$ is dense in $H_0^1[\alpha, \beta]$ (see, e.g. [1,3]), we can choose a function $\xi \in S$ with support in (α, β) such that

$$\int_{\alpha}^{\beta} (v'_1 - \xi')^2 < \varepsilon^2.$$

Then

$$\begin{aligned} |(v_1 - \xi)(x)| &= \int_{\alpha}^x (v'_1 - \xi') \leq \sqrt{\beta - \alpha} \sqrt{\int_{\alpha}^{\beta} (v'_1 - \xi')^2} \\ &< \varepsilon \sqrt{\beta - \alpha}. \end{aligned}$$

Thus, using the local boundedness of p and the local integrability of q , we get

$$\int_{\alpha}^{\beta} p(v'_1 - \xi')^2 + q(v_1 - \xi)^2 \leq \varepsilon^2 \left(\|p\|_{\infty, [\alpha, \beta]} + (\beta - \alpha) \int_{\alpha}^{\beta} |q| \right).$$

Thus we may choose a $v_2 \in S$ such that $\|v_2 - v_1\|_W < \varepsilon$.

Hence, $\|v - v_2\|_W < 2\varepsilon$ and

$$|a(u, v)| = |a(u, v - v_2)| \leq 2\varepsilon \|u\|_W.$$

The claim follows since ε is arbitrary. It follows from Eq. (18) that $v \mapsto a(u, v)$ is continuous on D_0 . Hence, $u \in \tilde{D} \subset D$. Furthermore, since $u \in W_2$, $\{z, u\}_a^b = \{u, z\}_a^b = 0 \forall z \in D_1$. Then $[u, z]_a^b = 0 \forall z \in D_1$. It follows from the characterization of the domains of self-adjoint extensions of L_0 in [16] that $u \in D_1$. Now since

$$a(u, v) = \langle L_1 u, v \rangle$$

for all $v \in D_0$ and since D_0 is dense in H , $L_1 u = 0$ and thus $u = 0$. ■

6. The Galerkin method for type I operators

Let $f \in H$. The discrete counterpart of Problem (V) reads: Find $u_m \in V_m$ such that

$$a(u_m, v) = \langle f, v \rangle \quad \forall v \in V_m. \quad (\text{Vm})$$

The following lemma is standard (see, e.g., [1]).

Lemma 18. *Problem (Vm) has a unique solution $u_m \in V_m$.*

Theorem 19. *The sequence $\{u_m\}$ of solutions of (Vm) converges to u (the solution of (V)) in the norm of W .*

Proof. The proof is a simple adaptation of Cea's Lemma to this setting. By (V) and (Vm)

$$a(u - u_m, v) = 0 \quad \forall v \in V_m.$$

Consequently, for any $v \in V_m$,

$$a(u - u_m, u - u_m) = a(u - u_m, u - v) \leq \sqrt{a(u - u_m, u - u_m)} \sqrt{a(u - v, u - v)}$$

which implies

$$\sqrt{a(u - u_m, u - u_m)} \leq \inf_{v \in V_m} \sqrt{a(u - v, u - v)}$$

and the latter term tends to 0 as $m \rightarrow \infty$ as a consequence of Proposition 17. ■

The above theorem and (8) give the following corollary.

Corollary 20. *The sequence $\{u_m\}$ converges to u in H .*

7. The Galerkin method for arbitrary self-adjoint operators

In this section we are given an arbitrary self-adjoint extension \widehat{L} of L_0 and are interested in how to implement the finite element-Galerkin method of Section 6 to solve approximately the equation

$$\widehat{L}u = f. \quad (19)$$

In the case $d = 0$, $\widehat{L} = L_0$ and no further consideration is needed. Thus we need only consider the case $d > 0$ here. For existence as well as uniqueness of the solution for (19) we assume throughout this section that 0 is not an eigenvalue of \widehat{L} . The equation

$$Lu = 0$$

has d solutions $u_1, \dots, u_d \in D$ which are linearly independent modulo D_0 .

Construct the functions ψ_1, \dots, ψ_{2d} in D such that, for $i = 1, \dots, d$, $\psi_i = u_i$ near a , $\psi_i = 0$ near b , $\psi_{d+i} = u_i$ near b and $\psi_{d+i} = 0$ near a (see Section 2). Using the boundary condition characterization of \widehat{L} discussed in Section 3,

we can construct d functions $\varphi_1, \dots, \varphi_d$ such that $\widehat{D} = D_0 + \text{span}\{\varphi_1, \dots, \varphi_d\}$. For example, if $d = 2$ and \widehat{L} is characterized by coupled boundary conditions as in Lemma 7, then, writing

$$\varphi_1 = \sum_{i=1}^4 \alpha_i \psi_i, \quad \varphi_2 = \sum_{i=1}^4 \beta_i \psi_i,$$

we have

$$\Phi_j(a) = e^{i\gamma} C \Phi_j(b), \quad j = 1, 2,$$

where

$$\Phi_j(a) = \begin{bmatrix} [\varphi_j, \psi_1](a) \\ [\varphi_j, \psi_2](a) \end{bmatrix}, \quad \Phi_j(b) = \begin{bmatrix} [\varphi_j, \psi_1](b) \\ [\varphi_j, \psi_2](b) \end{bmatrix}.$$

Substituting the expressions for φ_1, φ_2 we obtain (assuming that $[u_1, u_2](a) = [u_1, u_2](b) = 1$)

$$\begin{bmatrix} -\alpha_2 \\ \alpha_1 \end{bmatrix} = e^{i\gamma} C \begin{bmatrix} -\alpha_4 \\ \alpha_3 \end{bmatrix}$$

and a similar equation for the β 's. Thus taking two linearly independent vectors $\begin{bmatrix} -\alpha_4 \\ \alpha_3 \end{bmatrix}, \begin{bmatrix} -\beta_4 \\ \beta_3 \end{bmatrix}$, φ_1, φ_2 are determined. Next, the solution u of (19) can be uniquely written in the form

$$u = u_0 + r_1 \varphi_1 + \dots + r_d \varphi_d$$

with $u_0 \in D_0$. Thus,

$$\widehat{L}u = L_0 u_0 + r_1 \widehat{L}\varphi_1 + \dots + r_d \widehat{L}\varphi_d = f. \quad (20)$$

Denoting by P the orthogonal projection onto $\ker L = R_0^\perp$, and applying P to the above equation we get

$$P(r_1 \widehat{L}\varphi_1 + \dots + r_d \widehat{L}\varphi_d) = Pf. \quad (21)$$

Notice that P is given by

$$\langle Pv, u_i \rangle = \langle v, u_i \rangle, \quad i = 1, \dots, d.$$

Therefore, r_1, \dots, r_d can be determined by taking the inner product with $u_i, i = 1, \dots, d$ on both sides of (21) and solving the system

$$\begin{bmatrix} \langle \widehat{L}\varphi_1, u_1 \rangle & \dots & \langle \widehat{L}\varphi_1, u_d \rangle \\ \vdots & \ddots & \vdots \\ \langle \widehat{L}\varphi_d, u_1 \rangle & \dots & \langle \widehat{L}\varphi_d, u_d \rangle \end{bmatrix} \begin{bmatrix} r_1 \\ \vdots \\ r_d \end{bmatrix} = \begin{bmatrix} \langle f, u_1 \rangle \\ \vdots \\ \langle f, u_d \rangle \end{bmatrix}. \quad (22)$$

The following lemma tells us that this system is solvable.

Lemma 21. *The matrix of coefficients of the system (22) is nonsingular.*

Proof. Suppose not. Then a non-trivial linear combination of the rows is zero:

$$\sum_{i=1}^d \alpha_i \langle \widehat{L}\varphi_i, u_j \rangle = 0, \quad j = 1, \dots, d.$$

Let $v = \sum_{i=1}^d \alpha_i \varphi_i$. Then

$$\langle \widehat{L}v, u_j \rangle = 0, \quad j = 1, \dots, d.$$

Therefore, $\widehat{L}v \in R_0$. Since \widehat{L} is invertible by assumption, $\alpha_1 \varphi_1 + \dots + \alpha_d \varphi_d \in D_0$. This contradicts the linear independence of $\varphi_1, \dots, \varphi_d$ modulo D_0 . ■

Hence, Eq. (20) can be rewritten as

$$L_0 u_0 = f - r_1 \widehat{L} \varphi_1 - \cdots - r_d \widehat{L} \varphi_d := f_0. \quad (23)$$

Observe that, the solution of Eq. (23) can be obtained by solving the equation

$$L_1 u = f_0$$

where L_1 is any type I extension of L_0 . This can be seen as follows. Let v be a solution of the above equation. Then, since L_1 is an extension of L_0 ,

$$L_1 v = f_0 = L_0 u_0 = L_1 u_0.$$

Since L_1 is one-to-one, $u_0 = v$. It follows from this discussion and Theorem 19 that u_0 can be approximated in the norm of W by elements of S .

The finite element-Galerkin method can now be used as in Section 6 to compute u_0 . The algorithm discussed above can be summarized as follows.

1. Start with d linearly independent functions u_1, \dots, u_d in the kernel of L .
2. Construct the functions ψ_1, \dots, ψ_{2d} .
3. Given a self-adjoint extension \widehat{L} of L_0 , construct functions $\varphi_1, \dots, \varphi_d$ that characterize its domain.
4. Solve Eq. (22) to find r_1, \dots, r_d .
5. Solve approximately Eq. (23) by the finite element method to find u_0 .
6. Build the solution $u = u_0 + r_1 \varphi_1 + \cdots + r_d \varphi_d$ of (19).

We next show how the functions u_1, \dots, u_d in step 1 can be approximated (if not already known) with the help of an arbitrary type I operator L_1 with domain D_1 . The theoretical basis is as follows. Start with d regularizing functions $v_1, \dots, v_d \in D$ which are linearly independent modulo D_0 . Such functions are mostly found by inspection. From these functions we construct ψ_1, \dots, ψ_{2d} as explained in Section 2. Since D_1 is a d -dimensional extension of D_0 , the set ψ_1, \dots, ψ_{2d} can be replaced by a set $\varphi_1, \dots, \varphi_{2d}$ of linear combinations of them such that $\varphi_1, \dots, \varphi_d \in D_1$ and $\varphi_{d+1}, \dots, \varphi_{2d}$ are linearly independent modulo D_1 . The functions u_1, \dots, u_d are found as

$$u_i = \xi_i - \varphi_{d+i}, \quad i = 1, \dots, d,$$

where ξ_i is the unique solution in D_1 of the equation

$$L_1 u = L \varphi_{d+i}, \quad i = 1, \dots, d. \quad (24)$$

(By Theorem 19 the solution ξ_i can be approximated in the norm of W by elements of S .) Clearly, $Lu_i = 0, i = 1, \dots, d$. Furthermore, u_1, \dots, u_d are linearly independent modulo D_1 and, consequently, modulo D_0 . We can then use the finite element-Galerkin method of Section 6 to approximate the unique solution(s) of the equation(s) (24).

An alternative algorithm which avoids computing a function $u_0 \in D_0$ but still uses an arbitrary type I operator L_1 with domain D_1 can be derived as follows. Begin by finding a particular solution $u_p \in D_1$ for the equation

$$L_1 u = f. \quad (25)$$

(By Theorem 19 the solution u_p can be approximated in the norm of W by elements of S .) Now the solution \widehat{u} of (19) can be written as

$$\widehat{u} = u_p + r_1 u_1 + \cdots + r_d u_d$$

and it remains to determine r_1, \dots, r_d . This can be done by applying the boundary conditions characterizing \widehat{L} . For example, if $d = 2$ and \widehat{D} is described by coupled boundary conditions of the form

$$\widehat{U}(a) = C \widehat{U}(b),$$

then r_1, r_2 satisfy the system

$$\begin{bmatrix} [U_1(a) - CU_1(b)] & [U_2(a) - CU_2(b)] \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = CU_p(b) - U_p(a). \quad (26)$$

Lemma 22. System (26) determines r_1 and r_2 uniquely.

Proof. We need to establish that the matrix of coefficients of the system (26) is nonsingular. If this were not the case then there would exist a scalar m such that

$$U_1(b) - CU_1(a) = m(U_2(b) - CU_2(a)).$$

Rearranging, we get

$$U_1(b) - mU_2(b) = C(U_1(a) - mU_2(a)).$$

Therefore, $u_1 - mu_2 \in \widehat{D}$. Then $\widehat{L}(u_1 - mu_2) = L(u_1 - mu_2) = 0$. This contradicts the invertibility assumption of \widehat{L} . ■

Again the solution of (25) can be approximated by the finite element-Galerkin method of Section 6.

The alternative algorithm for computing the solution of (19) can now be summarized as follows.

1. If a basis u_1, \dots, u_d of $\ker L$ is known skip to step 6, otherwise, start with any functions $v_1, \dots, v_d \in D$ which are linearly independent modulo D_0 .
2. From these construct the functions ψ_1, \dots, ψ_{2d} .
3. From these construct the functions $\varphi_1, \dots, \varphi_{2d}$.
4. Solve the equation(s) (24) by the finite element method to approximate ξ_i , $i = 1, \dots, d$.
5. Set $u_i = \xi_i - \varphi_{d+i}$, $i = 1, \dots, d$.
6. Solve Eq. (25) by the finite element method to approximate u_p .
7. Compute r_1, \dots, r_d from the boundary condition characterization of \widehat{L} .
8. Build the solution $u = u_p + r_1u_1 + \dots + r_du_d$ of (19).

8. Numerical examples

In this section we give some numerical examples to illustrate the foregoing discussion and assess the numerical accuracy of the method. Before going into the examples we give some details about some numerical considerations and constructions.

We begin with the construction of the functions ψ_1, \dots, ψ_{2d} . Let us assume that we have a function $u \in D$ from which we want to construct a function $\psi \in D$ which equals u near a and 0 near b . The most general way to construct such a function is to take a suitable interval $[\alpha, \beta] \subset I$ and solve the initial value problem

$$\begin{aligned} \ell y &= f; \\ y(\alpha) &= u(\alpha), \quad y^{[1]}(\alpha) = u^{[1]}(\alpha), \end{aligned}$$

where f can be taken as a linear combination of the functions θ, ϕ of Section 2. This linear combination can be adjusted such that the solution y satisfies the conditions $y(\beta) = y^{[1]}(\beta) = 0$. ψ is then taken to be u on $(a, \alpha]$, y on $[\alpha, \beta]$ and 0 on $[\beta, b)$. See [10,16,20] for more details. Alternatively, and this is the approach we took, if the coefficient function p is C^1 on an interval $[\alpha, \beta] \subset I$ and $p(\alpha) \neq 0$, then we construct a third degree polynomial τ satisfying the four conditions $\tau(\alpha) = u(\alpha)$, $\tau'(\alpha) = u'(\alpha)$, $\tau(\beta) = \tau'(\beta) = 0$ and define ψ to be u on $(a, \alpha]$, τ on $[\alpha, \beta]$ and 0 on $[\beta, b)$.

Second, the integrals needed in (22) are computed by using the trapezoidal rule (on a finite interval). The trapezoidal rule is enough since it is $O(h^2)$ and we know that the finite element method with piecewise linear basis functions is $O(h^2)$ for regular problems.

Third, the derivatives needed to evaluate the expression ℓ are computed by a central difference method which is also $O(h^2)$ in the regular case.

Finally, the solution of the resulting linear tridiagonal system was found by using the MATLAB backward slash operator, which utilizes the LU factorization approach.

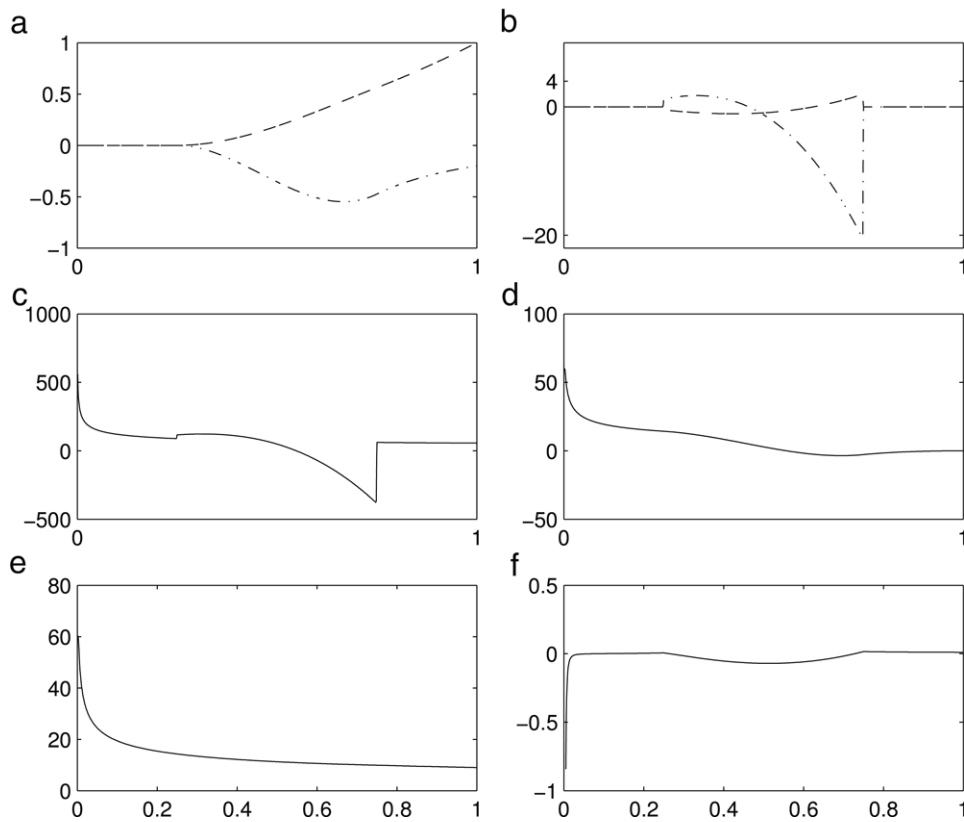


Fig. 1. Some results for Example 1: (a) $\psi_1(-), \psi_2(-)$; (b) $L\psi_1, L\psi_2$; (c) $f - rL\phi$; (d) the computed solution u_0 , and (e) the computed solution u (f) The pointwise errors.

Example 1. In this example we consider the Cauchy–Euler operator $\ell u = (-x^2 u')' + 6u$ on the interval $I = (0, 1)$. The equation $\ell u = 0$ has the two solutions $\theta = x^2$, $\phi = -\frac{1}{5}x^{-3}$ (taken so that $[\theta, \phi] = 1$). 0 is an LP and 1 is a regular point. Thus $d = 1$. We considered the problem

$$\widehat{L}u = 56x^{-1/3},$$

where \widehat{L} is the self-adjoint operator determined by the boundary conditions (11) (here a and b are exchanged) with $\tan \gamma = -\frac{1}{5}$. This value of γ was chosen so that $u = 9x^{-1/3}$ is the solution of the above equation. We ran the program with $a_N = .001$, $b_N = 1$ and $n = 1000$. A discretized form of the H norm of the difference between the computed and exact solutions was computed. The relative error in this case came out to be 1.3% over the whole interval. On the subinterval $[.01, 1]$, however, the relative error is only 0.6%. The difference is attributed to the fact that the solution u is infinite at 0. Fig. 1 shows some results for this example. Notice that the computed solution u_0 (figure (d)) satisfies the boundary conditions $u(1) = u^{[1]}(1) = 0$ at the regular endpoint. The boundary condition $(u^{[1]}u)(0) = 0$ at the singular endpoint is not immediately evident from the graph. It was verified computationally by comparing the values of the numerical approximations of the quantity $(u^{[1]}u)(x)$ with the true value $-3x^{1/3}$ near zero. This can also be inferred from the graph since u_0 and the solution u are identical near a and the computed solution is close to u .

Example 2. In this example we consider again the Cauchy–Euler operator of Example 1 on the interval $(0, \infty)$. In this case both endpoints are LP ($d = 0$). Therefore, the only self-adjoint operator in this case is L_0 itself. The function $u = \frac{1}{(1+x)}$ is the solution for the equation $L_0 u = f$, where $f = (6x^2 + 14x + 6) / (1+x)^3$. We ran the program with $a_N = .01$, $b_N = 100, 200, 300$ and $n = 1000, 2000, 3000$, respectively. The L^2 relative error reduced from about 6% for the first case to a little less than 4% in the last case. Fig. 2 shows the computed solution and the pointwise errors for the last case in the interval $[.01, 10]$.

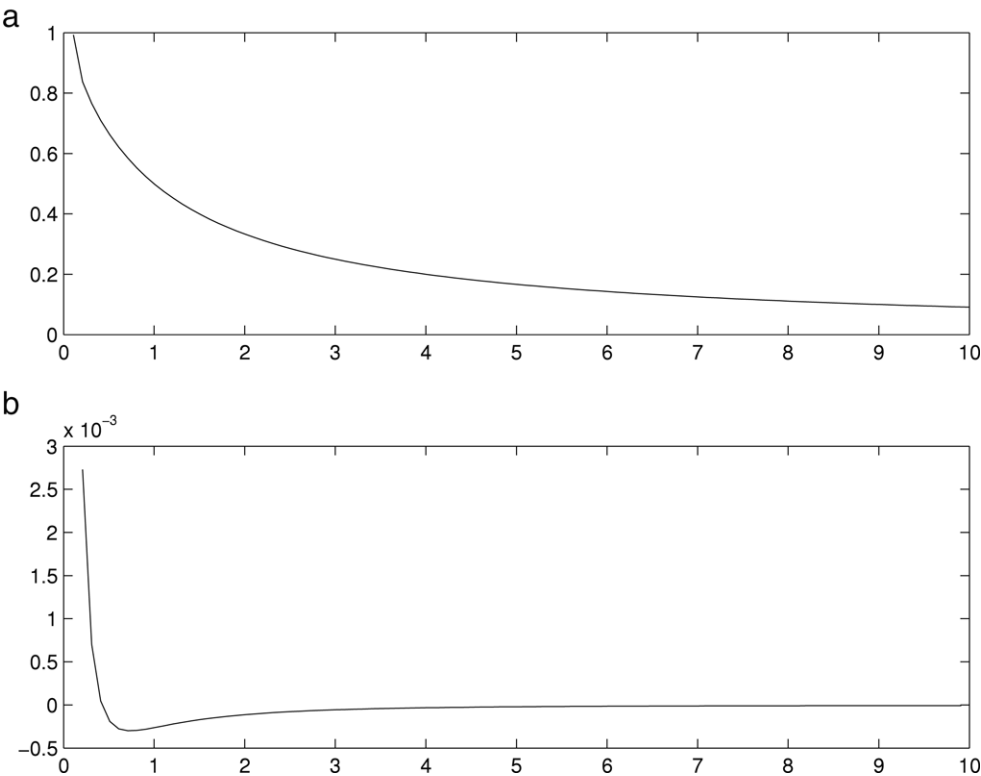


Fig. 2. Output for Example 2. (a) the computed solution, (b) the pointwise errors.

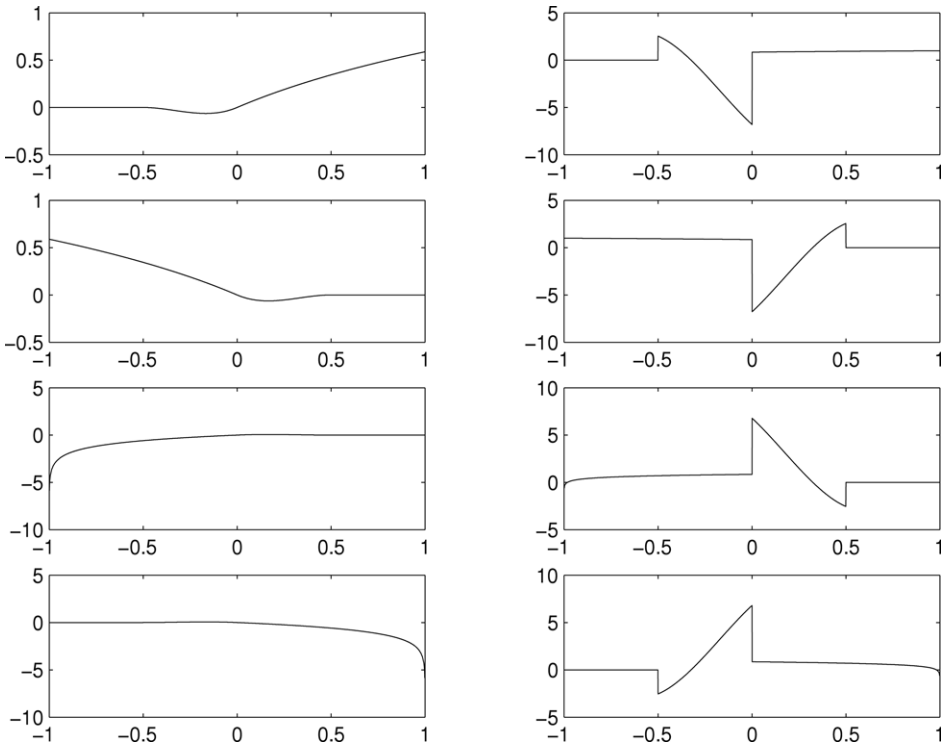


Fig. 3. The functions $\psi_1 - \psi_4$ (left column) and their images under L (right column).

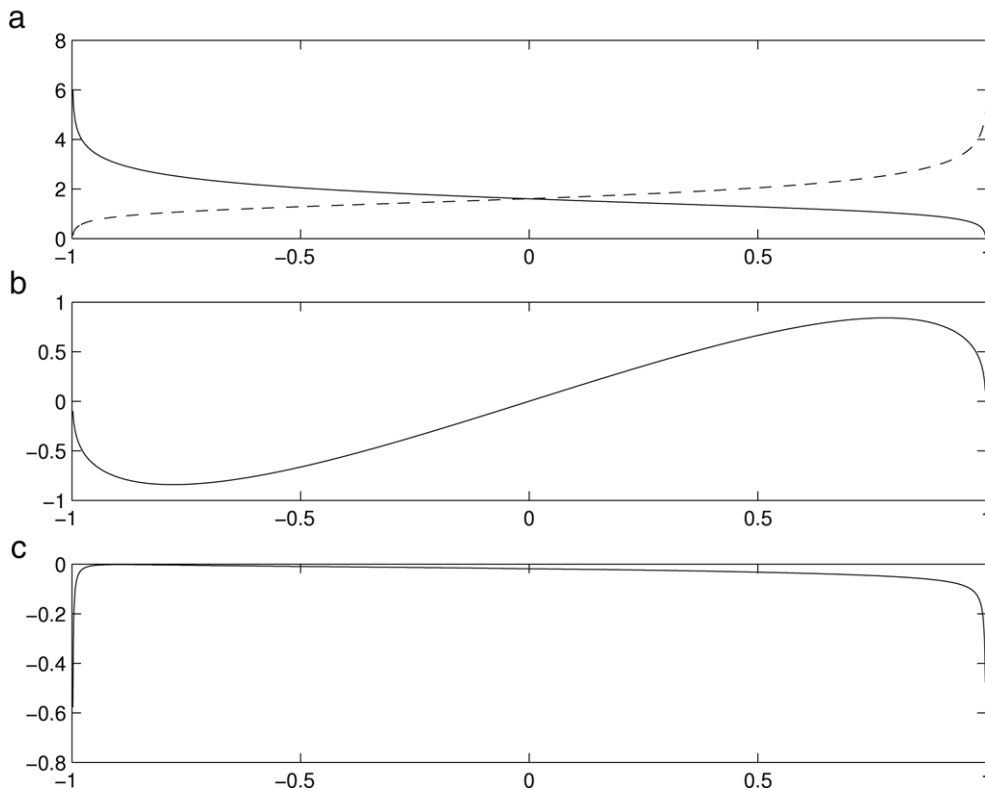


Fig. 4. Outputs for Example 3: (a) the functions u_1 (dashed), u_2 (solid); (b) the function u_p ; (c) pointwise errors.

Example 3. In this example we consider the Legendre operator $\ell u = -((1-t^2)u')' + \frac{1}{4}u$ on the interval $(-1, 1)$. The two functions $v_1 = \sigma \log(1+t)$, $v_2 = \sigma \log(1-t)$, where $\sigma = (2 \log 2)^{-1/2}$ are in D and are linearly independent modulo D_0 but they are not solutions of $Lu = 0$. Both endpoints are LC, therefore, $d = 2$. The function $u = t \log(1-t^2)$ is the solution of (19), where $f = \frac{9}{4} \log(1-t^2) + 6t^2$ and the domain of \hat{L} is determined by the coupled boundary conditions $U(a) = -U(b)$. We ran the second algorithm for this problem with $a_N = -.999$, $b_N = .999$, $n = 2000$. We obtained a relative L^2 error of about 4.6% on the whole interval. The error is about 2.3% on the interval $[-0.9, 0.9]$. Fig. 3 shows the functions $\psi_1 - \psi_4$ and their images under L for this example.

In this example, the functions $\psi_1, \psi_2 \in D_1$ while ψ_1, ψ_2 are linearly independent modulo D_0 . The next figure, Fig. 4 shows the functions u_1, u_2 , the particular solution u_p and the pointwise errors.

Acknowledgements

We would like to thank the referee for many suggestions which led to improvements in the paper. Research of the first two authors has been funded by King Fahd University of Petroleum and Minerals under Project number MS/Singular ODE/274.

References

- [1] K. Atkinson, W. Han, Theoretical Numerical Analysis: A Functional Analysis Framework, in: Texts in Applied Mathematics, vol. 39, Springer, New York, 2001.
- [2] H. Behnke, U. Mertins, M. Plum, Wiener's Eigenvalue inclusions via domain decomposition, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 456 (2003) 2717–2730.
- [3] C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, Texts in Applied Mathematics, Springer-Verlag, New York, 1994.
- [4] E. Davies, M. Plum, Spectral pollution, IMA J. Numer. Anal. 24 (3) (2004) 417–438.

- [5] K.O. Friedrichs, Spektraltheorie halbbeschränkter Operatoren, Math. Ann. 109 (1934) 465–487 and 685–713; also 110 (1935), 777–779.
- [6] M.A. El-Gebeily, A variational formulation for regular and singular self-adjoint differential equations, Ann. Differential Equations 18 (2002) 40–50.
- [7] M.A. El-Gebeily, Weak formulation of singular differential expressions in spaces of functions with minimal derivatives, Abstract and Applied Analysis 2005 (7) (2005) 691–705.
- [8] M.A. El-Gebeily, K.M. Furati, Real self-adjoint Sturm–Liouville problems, Appl. Anal. 83 (4) (2004) 377–387.
- [9] M.I. Gorbachuk, V.I. Gorbachuk, M.G. Krein's, Lectures on entire operators, in: Operator Theory, Advances and Applications, vol. 97, Birkhäuser Verlag, 1997.
- [10] A.M. Krall, A. Zettl, Singular self-adjoint Sturm–Liouville problems, Differential Integral Equations 1 (4) (1988) 423–432.
- [11] T. Kato, Perturbation Theory for Linear Operators, Verlag Springer, Berlin, Heidelberg, 1995.
- [12] J. Lahmann, M. Plum, On the spectrum of the Orr–Sommerfeld equation on the semiaxis, Math. Nachr. 216 (2000) 145–153.
- [13] M. Marletta, A. Zettl, The Friedrichs extension of singular differential operators, J. Differential Equations 160 (2001) 404–421.
- [14] U. Mertins, Asymptotic error-estimates for Rayleigh–Ritz-Approximations of self-adjoint eigenvalue problems, Numer. Math. 63 (2) (1992) 227–241.
- [15] U. Mertins, On the convergence of the Rayleigh–Ritz method for eigenvalue problems, Numer. Math. 59 (7) (1991) 667–682.
- [16] M.A. Naimark, Linear Differential Operators, Part II, Ungar, New York, 1968.
- [17] H.-D. Niessen, A. Zettl, The Friedrichs extension of regular ordinary differential operators, Proc. Roy. Soc. Edinburgh Sect. A 114 (1990) 229–236.
- [18] M. Plum, Bounds for eigenvalues of 2nd-order elliptic differential-operators, Z. Angew. Math. Phys. 42 (6) (1991) 848–863.
- [19] J. Pryce, Numerical Solution of Sturm–Liouville Problems, Clarendon Press, Oxford, 1993.
- [20] J. Weidmann, Spectral Theory of Ordinary Differential Operators, in: Lecture Notes in Mathematics, vol. 1258, Springer, Heidelberg, 1987.
- [21] A. Zettl, Sturm–Liouville problems, in: D. Hinton, P. Schaefer (Eds.), Pure and Applied Mathematics, Marcel Dekker, New York, 1997, pp. 1–104.
- [22] S. Zimmermann, Comparison of errors in upper and lower bounds to eigenvalues of self-adjoint eigenvalue problems, Numer. Funct. Anal. Optim. 15 (7-8) (1994) 943–960.